

**Evaluating the Psychometric Properties of the Appearance-related Repetitive
Behaviors Online Self-Report Scale**

Abdullah Mahimud & Anders Moberg

Institutionen för beteende- social- och rättsvetenskap (BSR), Örebro Universitet

PS303A - Examensuppsats

John Barnes, Ph.D.

2023-05-15

Abstract

Introduction: Body dysmorphic disorder (BDD) is an obsessive compulsive disorder that is marked by a distressing preoccupation with perceived physical flaws causing the individual to engage in compulsive behavior. The Appearance-related Repetitive Behavior Online self-report scale (ARBO) was developed to easily screen for appearance related behaviors believed to be common among BDD patients. Examining the psychometric properties of a newly developed scale is an essential part of its development. The aim of this study is to examine the factor structure, internal consistency and validity of the ARBO.

Method: Two studies were conducted in order to examine the psychometric properties of the ARBO. The first analysis was an Exploratory Factor Analysis (EFA) with the purpose of establishing a primary factor structure and the second consisted of a Confirmatory Factor Analysis (CFA) aimed at confirming the factor structure. Internal consistency tests were conducted in both studies. The validity of the factor model was examined after the CFA.

Result: The EFA revealed a two-factor structure that was later confirmed by the CFA after omitting items that were not consistent with the model. The internal consistency was exceptional in both samples. The confirmed factor model showed good convergent and divergent validity.

Discussion: Three items were removed during the EFA & CFA analyses in order to improve the model fit and internal consistency. The two factors are thought to be repetitive behaviors and reassurance seeking in an online environment. Screening for appearance-related repetitive online behaviors could be helpful in identifying patients with potential appearance concerns.

Table of Content

Abstract.....	2
Table of Content.....	3
Evaluating the Psychometric Properties of the Appearance-related Repetitive Behaviors Online Self-Report Scale.....	4
Insight and Trouble Diagnosing.....	6
Problematic Internet Usage, Social Media and BDD.....	8
The Appearance-related Repetitive Behavior Online Self-report Scale (ARBO).....	9
Psychometric Properties.....	13
Study Aim.....	17
Method 1.....	18
Survey development.....	18
Ethical Considerations.....	19
Procedure.....	19
Statistical Analysis.....	22
Result 1.....	23
Exploratory Factor Analysis.....	23
Internal Consistency & Intercorrelations Between Factors 1.....	25
Brief Discussion.....	26
Method 2.....	26
Measurements.....	27
Convergent validity.....	28
Divergent validity.....	29
Ethical Considerations.....	30
Procedure.....	30
Statistical Analysis.....	31
Result 2.....	32
CFA results.....	32
Validity.....	34
Internal Consistency & Intercorrelations Between Factors 2.....	34
Discussion.....	35
Factor analysis.....	36
Convergent & divergent validity.....	40
Limitations & Strengths.....	43
Future Research.....	46
Acknowledgment.....	48
References.....	49

Evaluating the Psychometric Properties of the Appearance-related Repetitive Behaviors Online Self-Report Scale

Body dysmorphic disorder or BDD is an obsessive compulsive disorder characterized by a preoccupation with perceived defects in an individual's physical appearance (Phillips, 2009). These preoccupations cause obsessions and engagement in compulsive repetitive behaviors. Due to these obsessions and compulsions, the individual often experiences a deep sense of distress or dysfunction, or both, in everyday life (APA, 2013). Common areas for these perceived flaws to manifest are: 1) individual's skin, such as acne, wrinkles, skin tone and/or color; 2) nose, such as shape or size; 3) hair, such as body hair and thinning of the hair; 4) teeth, believing them to be crooked or discolored; 5) the stomach, often leading individuals to believe they are obese; and 6) muscles, especially for males, leading the individual to believe that they are too skinny (Phillips, 2009; APA, 2013; Schieber, 2015; Möllman et al., 2017; Bridley & Daffin, 2022).

BDD often causes impairment in different areas of life with more than half of patients affected by BDD reporting an impaired social life and many reporting having trouble finding a romantic partner (APA, 2013; Brohede et al., 2015). A further 41% of patients with BDD have self-reported impairment in regards to their occupational life with unemployment and sick-leave being far more common relative to the general population (Brohede et al., 2015). Beyond social and occupational impairment, BDD patients are at an increased risk of developing comorbid disorders (Phillips, 2009). Common comorbidities for BDD patients are Social Anxiety Disorder (SAD), Obsessive Compulsive Disorder (OCD) and Major Depressive Disorder (MDD) (Phillips, 2009; APA, 2013; Brohede et al., 2015). Lastly, patients with BDD often suffer from suicidal ideations and self harm (Phillips, 2009; APA, 2013). A recent study by Oshana et al. (2020) noted that BDD patients were four times more likely to report suicidal ideations compared to the US national average. Other studies report

that 30-36% of all BDD patients have experienced suicidal ideations (Schiber, 2015; Möllman 2017).

Individuals with BDD often report spending 3 to 8 hours every day obsessing over their perceived flaw and engaging in compulsive behaviors. (APA, 2013). The BDD obsessions often result in mental acts such as fixating on a specific body part and comparing one's appearance with others (Phillips et al., 2006; Phillips, 2009 & APA, 2013). These mental acts are experienced as distressing and difficult to control and can often limit the individual in a variety of contexts such as occupational and social etc (Phillips, 2009 & APA, 2013). Additionally the individual may engage in compulsions, often as an attempt to fix the perceived flaw in their appearance. These fixes can include anything from excessive grooming and skin picking to seeking dermatological help or plastic surgery (Crerand, 2005; APA, 2013; Veale et al., 2016; Möllman et al., 2017). Further, BDD patients often engage in behaviors such as excessive tanning, changing clothes repeatedly throughout the day and compulsively buying make-up and other beauty products (Bridley & Daffin, 2022). Furthermore, individuals with BDD often excessively pick at their skin, causing damage and there are even cases where individuals have attempted to perform surgery on themselves (APA, 2013).

BDD is not only a severe and debilitating disorder but it is also a relatively common one. The prevalence of BDD is roughly 2.4% among adults, it being slightly more common among women (2.5%) than men (2.2%) (APA, 2013). For example, Brohede et al. (2015) found the prevalence of BDD in Swedish females to be 2.1%. They also found the prevalence of BDD to depend on the individual's age, with a prevalence of 3.4% among 18-30 year olds compared to 1.3% among 31-45 year olds and 1.6% among 46-60 year olds respectively. Thus, BDD is both distressing and common, often affecting younger people and with

prevalence rates exceeding more known disorders such as Obsessive-Compulsive Disorder (OCD) (APA, 2013).

Insight and Trouble Diagnosing

A common aspect of BDD is that patients often lack insight regarding the psychological aspects of their disorder (Buchanan et al., 2011; APA, 2013). This lack of insight can in some cases be so severe that it is considered delusional, with patients experiencing very low to no insight into the cognitive aspects of the disorder (Phillips, 2004a). Additionally, many who suffer from BDD are reluctant to even consider the possibility that their problem is psychological. Eisen et al. (2004) found that reluctant patients were more likely to believe that others shared their perspective but were prone to disagree when others' offered positive feedback about their appearance. BDD sufferers tend to believe that their flaws are real and that others share their view (Phillips, 2004b; Phillips, 2009) and this lack of insight causes many to act drastically, by seeking treatment from dermatologists, dentists or plastic surgeons (Crerand, 2005).

Many victims of BDD disregard the psychological aspects of their disorder and seek non-psychiatric care. Crerand et al. (2005) observed that 71% of patients with BDD had at some point during their illness sought non-psychiatric treatment for their perceived flaws with 64% receiving such treatment, most commonly from a dermatologist. Similarly, Veale et al. (2016) observed high rates of BDD patients in several non-psychiatric treatment fields with 13.2% in general cosmetic surgery, 20.1% in rhinoplasty surgery, 11.2% in orthognathic surgery, 5.2% in orthodontics/cosmetic dentistry settings, 11.3% in dermatology outpatients, 9.2% in cosmetic dermatology outpatient clinics and 11.1% in acne dermatology clinics. However, few patients experienced a reduction of BDD symptoms. Among the receivers of non-psychiatric care only 17.7% of patients reported an improvement regarding the perceived

flaw and as few as 3.6% experienced a significant reduction of BDD symptoms. (Crerand et al., 2005).

Apart from the difficulties BDD patients themselves have with recognizing the psychological aspects of their problems clinicians also face difficulties identifying and diagnosing BDD. Recent studies show that only 27% of patients later diagnosed with BDD were referred to clinics with concerns about appearance anxiety. Most patients were instead referred for concerns about similar disorders such as OCD (Buckley et al., 2019).

Additionally, difficulties in diagnosing BDD may also stem from patients being unwilling to disclose information about their problems. Singh, & Veale (2019) found that BDD patients are often unwilling to disclose their diagnosis to a therapist even when in a psychiatric or therapeutic setting. The unwillingness is reportedly often due to the patient feeling ashamed about their disorder or being afraid that their therapist will misunderstand them or be judgmental (Singh, & Veale, 2019).

Due to patients' limited insight and the difficulties in identifying BDD, new methods that assist clinicians are needed. Patients who suffer due to BDD need correct and helpful treatment (Phillips, 2009). Further, patients who are not aware of their disorder need assistance in gaining insight to prevent them from seeking unhelpful treatment options (Veale et al. 2016). Holland & Tiggemann, (2016) have previously theorized that people's internet usage is a potential way to screen for BDD. Specifically by examining appearance-related activities on Social media, otherwise known as Social Networking Sites (SNS) and Problematic Internet Usage (PIU). Given that over 90% of people in Sweden use the internet on a daily basis (Internetstiftelsen, 2021) this might be a potential domain in which to better screen for BDD sufferers.

Problematic Internet Usage, Social Media and BDD

Internet usage has steadily increased over the last few decades. In Sweden the non-profit organization managing the largest local domain in the country, .se, known as Internetstiftelsen (2021) estimates that 94% of the population uses the internet on a daily basis. Further, the use of social media such as Facebook, Instagram and Snapchat remains high with 94% of Swedes over the age of 8 reportedly having used an SNS during the past year. This high amount of internet and SNS usage is not without its potential risks. Due to the internet becoming a staple of people's daily life, vulnerable individuals may experience PIU (Shapira et al., 2000). PIU can generally be defined as when internet usage becomes excessive and causes significant distress and functional impairment in a variety of contexts. Additionally, individuals who engage in PIU often develop comorbidities such as depressive symptoms and excessive repetitive behaviors online (Shapira et al., 2000). PIU has also been shown to correlate with a more negative body image and may lead to neglect of financial, social and occupational commitments (Koronczi et al., 2013).

Apart from PIU and the negative body image it correlates with, there is also a connection between appearance comparison on SNS's, and body dissatisfaction (Holland & Tiggemann, 2016; Rousseau, 2017; Ryding & Kuss 2020). There are signs that SNS usage has become a way for especially younger people to engage in appearance comparison with each other. Rousseau (2017) found that SNS usage positively predicted appearance comparison and body dissatisfaction among American female adolescents. Further, a systematic review by Ryding & Kuss (2020) showed that excessive time spent engaging with an SNS increases tendencies for self-criticism and dysfunctional attitudes towards the self. Participants also tended to suffer from decreased self-esteem, lower body image, negative appearance evaluation and increased risk of developing BDD symptoms. This growing body of evidence connects PIU and certain aspects of SNS usage with body dissatisfaction and the

development of symptoms similar to that of BDD. Consequently, it is reasonable to assume that appearance-related repetitive behaviors might also manifest in an online environment. Identifying such repetitive behaviors would be beneficial for clinicians who face trouble recognizing appearance-related problems as well as the patients who are ashamed of their issues and/or lack insight into their problems.

In an attempt to increase the detection of BDD online the novel Appearance-related Repetitive Behaviors Online Self-Report Scale (ARBO) (Barnes et al., 2020) has been developed to assess individuals' appearance-related repetitive behaviors online. Preliminary assessment of the ARBO as an online screening tool is promising and ongoing development of the scale is currently in progress. In addition to other recently developed tools that assess repetitive appearance related behavior online (see Zimmer et al., 2021) development of the ARBO may facilitate improvement in the detection of BDD.

The Appearance-related Repetitive Behavior Online Self-report Scale (ARBO)

The ARBO aims at identifying individuals manifesting appearance-related behaviors in an online environment. The scale took inspiration from previously identified common behaviors that BDD patients engage in and modified these to fit into an online environment.

Specifically, items were based on repetitive behaviors common in BDD (Phillips et al., 2006; Phillips, 2017). The ARBO consists of 20 items. The full list of ARBO items can be seen in Table 1.

Table 1*ARBO Items*

Item number	Item
1.	Do you use social media?
2.	Estimated, how many pictures are there of you on social media?
3.	How often do you post pictures of yourself on social media?
4.	How often do you remove pictures or remove tags of yourself on social media because of distress related to your appearance concerns?
5.	How often do you post pictures of yourself on social media in order to get feedback regarding your appearance concerns? (For example, uploading pictures or write posts about appearance)
6.	How often do you review pictures of yourself on social media because of your appearance concerns?
7.	How often do you edit pictures of yourself to improve your appearance before you upload them to social media, because of distress related to your appearance concerns?
8.	When you upload pictures of yourself to the internet, how often do you try to hide parts of your appearance due to distress related to your appearance concerns? (For example, editing pictures, using filters or by using accessories or decals to hide parts of your appearance?)
9.	When you look at pictures of other people online and on social media, how often do you compare parts of your appearance with them?
10.	How often do you send pictures of yourself to friends or family to ask about the parts of your appearance which causes you distress?
11.	How often do you use the internet to search for information regarding the parts of your appearance which cause you distress?
12.	How often do you use the internet to search for information about cosmetic surgery or treatments for the parts of your appearance which cause you distress?
13.	How often do you use the internet to look at videos of cosmetic treatments, such as surgical procedures, related to your appearance concerns?
14.	How often do you use the internet for buying appearance related products due to your appearance concerns?
15.	How often do you use or visit web pages that purports to rate a person's attractiveness based on their appearance, to get feedback on your appearance concerns?
16.	How often do you use the internet, for example Google Search or social media, to find pictures or people to compare the parts of your appearance which cause you distress?
17.	How often do you avoid dating pages or dating apps, such as Tinder or Match.com, because of your appearance concerns?
18.	How much time, in total, do you spend on actions online and social media that are related to the parts of your appearance which you feel distress over? Estimate as best you can.
19.	How much distress do these actions cause you?
20.	How would you feel if you were prevented from doing these actions? How much distress would you feel? Estimate the degree of distress/frustration you would feel if these actions were interrupted.

Note. Item 1 is answered yes or no. Item 2 is answered from 0 to 100 pictures or more. Items 3 - 17 are answered on a 5-point Likert scale from never (1) to very often (5). Item 18 is an estimation of time spent on the above behaviors and items 19 and 20 are answered on a scale of 1-10.

The first item asks participants whether they use an SNS or not. This was done in order to ensure that low scores on the following questions were not due to participants simply not being users of an SNS. The second item asks participants how often they post pictures of themselves. The next 15 items gather information about how the repetitive behavior that characterize BDD patients manifest in an online environment. Items 3-9 are aimed at examining PIU through the use of SNS whilst items 9-15 examine PIU through general internet usage. These items are rated on a five-point Likert Scale, ranging from *Never* to *Very often* (1-5).

Beyond the initial items and the 15 symptom items the ARBO consists of three final items that measure time spent on repetitive behaviors and severity. The ARBO asks participants how much time in total is spent per day on online behavior related to their appearance concerns with answers being 1) *less than 1 hour*; 2) *1-3 hours*; 3) *3-8 hours* and; 4) *more than 8 hours*. The ARBO then concludes with 2 items that measures the severity of the manifested symptoms. These 2 items were rated on a scale from 1 to 10, ranging from *No problems at all* to *Very much*.

The 15 core items of the ARBO aim to capture how behaviors common in BDD manifests online. The participants are asked how often they post images of themselves online (*item 3*). The ARBO then asks how often the participant attempts to hide certain parts of their appearance. In the ARBO this is done by asking about deleting images or tags (*item 4*), or photo-editing (*item 7 & 8*) due to appearance concerns. The ARBO changes behaviors such as appearance comparisons in a real life environment to image searching (*item 16*) followed by appearance comparison (*item 9*). Reassurance and feedback seeking in BDD online is assessed by asking if participants excessively review images of themselves before posting (*item 6*), send images to friends and family to get reassured (*item 10*) or post images on SNS or appearance comparison sites (*item 5 & 15*). The ARBO also checks for confirmation bias

common to BDD by asking the participant if they search the internet for information about a perceived appearance-flaw (*item 11*). The ARBO also examines if an individual purchases appearance related products such as make-up online to hide appearance flaws (*item 14*). The ARBO also analyzes whether the participant is interested in cosmetic surgery. This is done by asking if the participant actively searches for information on plastic surgery (*item 12*) or if they often watch videos of plastic surgery (*item 13*). Finally, The ARBO asks if the participant avoids dating apps (*item 17*). Many BDD patients report difficulties in finding a romantic partner (Brohede et al., 2015) and this item aims to encapsulate this problem.

A preliminary analysis of the ARBO (Barnes et al., 2020) showed the ARBO to be reliable with a Cronbach's Alpha of .87 with item 1 excluded or .88 with items 2 and 3 also excluded. The ARBO exhibited good convergent validity with the Body Dysmorphic Disorder Questionnaire (BDDQ), a common diagnostic tool when screening for BDD (Phillips, 1995). Further, the severity measures of the ARBO were found to correlate well with the Body Dysmorphic Disorder-Symptom Scale (BDD-SS). The BDD-SS is an established scale measuring which BDD symptoms are prevalent and the severity of those symptoms (Wilhelm et al., 2015). Finally, the ARBO was found to not be collinear with established scales measuring BDD. This suggests that ARBO offers a unique perspective on repetitive behaviors common to BDD, specifically, in an online environment (Barnes et al., 2020).

Recently, Zimmer et al. (2021) developed a scale aimed at measuring BDD through social media - the Social Media Appearance Preoccupation Scale (SMAPS). Whilst similar, the ARBO does however differ from the SMAPS since it measures the level of distress through general online behavior and PIU, not just through activities on SNS. The ARBO also includes items measuring PIU intended to seek out information about cosmetic surgery and the purchasing of products intended to hide perceived appearance flaws, areas that are not

included in the SMAPS (Zimmer et al., 2021). As seen by Koronczi (2013) individuals who are dissatisfied with their appearance are more likely to engage in PIU. By not limiting the ARBO to only examine appearance-related behaviors through SNS usage it is able to detect a broader range online behaviors and consequently people who may be struggling with appearance-related repetitive behaviors online.

Preliminary assessment of the ARBO found by Barnes et al. (2020) is promising but further validation of the scale is required. The aspects of the ARBO which have not yet been analyzed are the psychometric properties and their implications.

Psychometric Properties

Psychometrics is a field in behavioral sciences concerned with quantifying psychological phenomena and developing measurement methods. A criteria of psychometric research is the replicability of measurement results obtained from a certain sample in order to generalize it. When developing a questionnaire or psychological test, the aim is to measure aspects of hypothetical non-directly observable constructs (AERA, APA, & NCME, 2014; DeVellis, 2017). Therefore, questionnaires are in essence measuring non-observable constructs which poses a problem for empirical replication. A second obstacle for replicability is the fact that each individual is unique. This uniqueness is reflected in responses on measurements such as the ARBO (Kyriazos, 2018). Consequently, to ensure replicability it is an essential to test the quality of newly developed psychological measurements. In psychometric research, this is done through various methods including testing the factor structure, reliability and validity of the newly structured scale.

Factor structure refers to identifying the number of underlying constructs that are reflected in a measurement scale (DeVellis, 2017). Exploring the factor structure reveals the underlying factors or latent variables of the scale. These two terms, factors and latent variables, will be used synonymously for the remainder of the thesis.

Exploring the factor structure is a necessary step in order to connect the theoretical structure of the scale with the questionnaires' items. Practically, the factor structure is explored by conducting a factor analysis to explore the number of latent variables in the questionnaire (Knehta et al., 2019). Latent variables are the result of multiple items clustered together to construct a new, non-observable factor (Field, 2018). When first produced, the number of potential latent variables are equal to the number of items prior to analysis (Clark & Watson, 1995). In order to create a sensible factor structure, multiple potential latent variables need to be excluded. This process is the second objective of factor analysis (Clark & Watson, 1995). The multidimensionality of a newly constructed scale needs to be simplified so it can easily identify the factors responsible for the variance in the results (Anastasi, 1985). In sum, factor analysis is performed to identify and examine the latent variables of a measurement scale. It is recommended by psychometricians and statisticians to begin the process of validating with conducting an exploratory factor analysis (Clark & Watson, 1995).

The psychometric property reliability refers to the extent that test scores are free from measurement errors. It is important to assess the extent to which a scale of measurement is consistent if the measurement is repeated multiple times in similar contexts (Boateng et al., 2018). There are multiple types of reliability that can be assessed when developing new scales. Which type of reliability to look at depends on the type of data researchers work with and the preliminary data on the ARBO permits the assessment of internal consistency, that is, the consistency of scales scores across multiple respondents (Dima, 2018).

There are multiple statistical tests designed to assess internal consistency. The prominent reliability test used in psychological research is called Cronbach's alpha (DeVellis, 2017). Alpha is a coefficient that measures the average correlation between a scales' items and average intercorrelation of the total score. Alpha coefficient assumes a normal

distribution of responses to questionnaire and an equal items' factor loadings. However, a large body of research suggests that Cronbach's alpha is not the optimal way of measuring internal consistency. The concern is that it lacks precision and overlooks a critical part of internal consistency which is the factor structure (DeVellis, 2017). Alpha's assumptions are rarely met in psychological research for example responses to psychometric scales usually have other distributions than the bell curve and unequal factor loadings (McNeish, 2018). Consequently, a more robust measurement method of internal consistency is needed. DeVellis (2017) mentions numerous alternatives to Alpha coefficient. An alternative that can examine internal consistency in a more precise and highly certain manner is Mccdonald's omega. Similar to alpha, coefficient omega uses the same rationale of variance of total scores (DeVellis, 2017). In contrast to Alpha coefficient, omega utilizes item loadings on common factors to assess internal consistency. Further, it is a more reliable measurement of internal consistency because, in contrast to alpha, it abandons the assumption stating that items' factor loadings are equal. Omega's reliance on a factor analytical approach makes it a more reliable and robust option for assessing scales' internal consistency (Trizano-Hermosilla & Alvarado, 2016).

Validity is also examined when developing a new measurement scale. There are multiple perspectives and conceptualizations on what validity is. The umbrella definition of all of these perspectives is how accurate the conclusion derived from a measurement scale is. Meaning, it examines how well a scale is evaluating the construct it intends to measure (Streiner et al., 2015; Kazdin, 2016; Boateng et al., 2018; Dima, 2018). Similarly to the different types of reliability, there are various ways to validate a measurement scale. Choosing which type of validity is also primarily influenced by the type of data a researcher has. The preliminary data on the ARBO permits only the examination convergent and divergent validity. Both convergent and divergent validity examine the association between a

new scale and already established ones but they have contrasting objectives. The former is interested in seeing the association between the new construct and constructs that are conceptually and empirically similar. The latter tries to demonstrate that there is a weak, non or negative correlation to constructs that are conceptually different (Boateng et al., 2018).

Preliminary analysis of the ARBO has shown that it possesses decent validity and reliability and appears to measure something akin to BDD (Barnes et al., 2020). However, without a factor analysis no conclusions can be drawn about what the ARBO actually measures. A factor analysis can also solve the potential problem of multicollinearity (Field, 2018). That is, if multiple items in the ARBO are found to be collinear then they can either be combined or one of them can be removed from the factor model. A factor analysis will also reveal item redundancy. It will show if the ARBO contains certain items that are not linked to the latent variables, allowing the researcher to delete them from the analysis (Joachim et al., 2016).

Preliminary assessment has shown that the ARBO correlates well with concurrent scales that measure BDD, such as the BDD-SS (Barnes et al., 2020). This means that the ARBO measures something similar but not exactly the same thing. A factor analysis is helpful in this area as well. Establishing a factor structure will cluster items together into latent variables. These latent variables can then be argued to theoretically represent certain aspects of appearance-related online behaviors or BDD based on the items they contain (Field, 2020). If several latent variables are found a factor analysis will help organize the different clusters of items. This creates a model where items can be organized into the latent variable where they fit the best. When this process is done correctly on a theoretically sound scale the latent variables will represent clear themes of the overarching concept that the scale aims to measure (Field, 2018; Knekta et al., 2019). In the ARBO's case this is appearance-related online behaviors which is linked to BDD (Holland & Tiggemann, 2016).

Thus, a clear argument for why the ARBO differs from similar scales that specifically aims to measure BDD can be made and the unique properties of the ARBO can be highlighted.

Apart from the factor structure, examining the psychometric properties of the ARBO and establishing a factor structure will make it more accurate. This increases the true positives it captures, making it more inline with the real world (Knehta et al., 2019). Further, validating the ARBO provides practitioners with an objective screening tool for measuring appearance-related online behaviors. A reliable and valid testing measurement aids the practitioners in making bias free decisions. It also enables them to draw more thoughtful conclusions. Having reliable and valid measurement increases the efficiency of the diagnostic process. This saves clinicians time and resources that can then be reallocated to other areas (Rust et al., 2020). This becomes more important as clinicians continue to face difficulties in identifying BDD (Crerand, 2005; Buckley et al., 2019; Singh, & Veale, 2019).

Study Aim

The aim of this study is to explore the psychometric properties of the ARBO self-report scale on two samples collected through snowball recruitment and consisting of mostly Swedish residents with a mean age of 33.5 and 29.1 respectively. The study will specifically be analyzing the factor structure, the internal consistency and convergent and divergent validity. Due to the nature of validating scales the thesis does not include a hypothesis and will constitute two separate studies. Study one aims to explore the latent variables of the ARBO self-report scale using an Exploratory Factor Analysis (EFA). Study two aims to confirm the explored factor structure in study one using Confirmatory Factor Analysis (CFA). Further aims with study two is to examine the internal consistency of the factor model and to examine whether the factor model is valid or not when measuring appearance-related behaviors in an online environment. In order to examine the validity of the factor model study two will look at divergent and convergent validity. The factor analysis, reliability analysis and the validity

analysis will all specifically be aimed at examining the ARBO, and later the factor structure and model, in the online environment it was designed for.

Method 1

The main recruitment approach involved Snowball recruitment through online advertisements via SNS groups during spring of 2020. Being a Swedish resident and older than 18 was listed as a requirement for participation in the survey. A total of 1501 participants answered the survey. 76% of the participants identified as women, 23% as men and 1% as others. Ages ranged from 18 to 85, ($M = 33.5$, $SD = 11.7$). The resulting sample consisted of 97.3% Swedish residents and 2.7% with *other* as their stated residency. For information about descriptives, see Table 3.

Due to the multicultural and diverse nature of the Swedish population, a decision was made by the authors to retain participants despite nationality or country of origin. The percentage of the participants who listed *other* as their country of origin was ~2.7% for sample 1 and ~11.5% for sample 2. This number does differ from the actual number of Swedish inhabitants born outside of Sweden which is ~20.4% (SCB, 2022). Considering how varied this group is, with people in it originating from every continent in the world, a decision was made to retain the participants who listed their country as *other* and disregard nationality. The decision to do so was aided by conducting a mean comparison analysis of the two groups. The results suggested that there is no material significance in the difference of ARBO means between people that choose *Sweden* and *other* as their nationality.

Survey development

The data used in study 1 was collected through a Google Forms survey. The questionnaire contained a variety of demographic questions about age, gender, citizenship, employment, and socioeconomic-status as well as about BDD symptoms and anxiety and depressive symptoms. The survey also asked about specific internet behaviors and thought processes.

Additionally, respondents were asked to report if they had taken a sick day due to worries about their appearance and whether they still do. Participants were also asked if they had ever sought treatment or therapy for these issues (cosmetic, psychological etc). The BDDQ was utilized to determine risk for BDD. An individual with no prior knowledge of the project reviewed the questionnaire in order to look for mistakes or sections that were difficult to comprehend before it was posted online. No changes were done after the review.

Ethical Considerations

The ethics of the original unpublished thesis on the ARBO by Barnes et al. (2020) was approved by Örebro University. The same committee later approved the ethics of this study before the datasets used in the study were handed out to the authors. No personal information was included in this study and all participation was voluntary. All participants were given written information on how their answers would be used according to etikprövningsmyndighetens guidelines (2021). Participants were also informed that they could withdraw their answers from the study. All data was stored on a secure server at Örebro university in order to preserve the security of the data. Results were reported so that no answers could be linked to the original participants.

Procedure

Prior to the analysis, items 1, 2, 18, 19 and 20 were excluded from the factor analysis. Item 1 and 2 asked whether the respondent uses social media and how many pictures of them that existed online. Item 1 was deleted due to the dichotomous nature of the item (*yes* or *no*). Participants who answered *no* would further have been labeled as missing and removed during analysis. Item 2 asked participants how many pictures the individual had uploaded to an SNS. This was considered to be too similar to item 3 which asked how often the individual uploaded images. Item 3 was deemed to be better formulated and better at encapsulating the intention of the question. Including both items in the analysis also risks creating a separate

latent variable for these two items. Since the minimum number of items for a latent variable in a factor analysis is 3 (Field, 2018) one had to be deleted. Item 3 was thus retained while item 2 was removed. Item 18 asked participants how many hours per day were spent on their repetitive behaviors but was removed as it was continuous and not categorical. Further, participants who spent no time on appearance-related repetitive behaviors were grouped with participants who spent an hour on the behavior. This was deemed to be too large of a span and further reason to delete the item. Finally, item 18 would potentially have to be sorted into its own latent variable, measuring duration; it was thus removed from the analysis for the same reason as item 2. Finally, items 19 and 20 asked participants about severity and distress caused by the symptom items on a scale of 1 to 10. These items were also removed as they would merit their own latent variables measuring severity.

The numbering of the remaining 15 items submitted to the analysis began from 1 (e.g. item 3 will become item 1, item 5 will become 3, item 15 will become 13..etc). The items measuring the characteristics of the repetitive behavior were then recorded from the original Swedish wording to a standardized, numerical Likert scale (*Never* = 1, *Seldom* = 2, *Sometimes* = 3, *Often* = 4 and *Very Often* = 5) (for Item 6 *very often* = 5 was changed to *always* = 5) Item 6 also included an answer stating *There are no pictures of me* and item 7 and 8 included the answer *I never upload pictures*. These answers were coded as 0 as they did not fulfill any certain niche compared to *never*. Finally, listwise deletion was used throughout the analysis to delete participants with missing data. Thus, participants who had not completed the entire survey were excluded from the analysis. A total of 1395 participants remained after deletion. No participants were removed from the survey due to residency or nationality. Descriptive statistics were checked in order to assess if mean, standard-deviation and variance were within reason and to see if any further missing data existed in the sample.

The full list of items with correct numbers post initial item deletion can be seen in Table 2. The participants' demographics can be seen in Table 3.

Table 2

ARBO Items

New Item Number	Item
1.	How often do you post pictures of yourself on social media?
2.	How often do you remove pictures or remove tags of yourself on social media because of distress related to your appearance concerns?
3.	How often do you post pictures of yourself on social media in order to get feedback regarding your appearance concerns? (For example, uploading pictures or write posts about appearance)
4.	How often do you review pictures of yourself on social media because of your appearance concerns?
5.	How often do you edit pictures of yourself to improve your appearance before you upload them to social media, because of distress related to your appearance concerns?
6.	When you upload pictures of yourself to the internet, how often do you try to hide parts of your appearance due to distress related to your appearance concerns? (For example, editing pictures, using filters or by using accessories or decals to hide parts of your appearance?)
7.	When you look at pictures of other people online and on social media, how often do you compare parts of your appearance with them?
8.	How often do you send pictures of yourself to friends or family to ask about the parts of your appearance which causes you distress?
9.	How often do you use the internet to search for information regarding the parts of your appearance which cause you distress?
10.	How often do you use the internet to search for information about cosmetic surgery or treatments for the parts of your appearance which cause you distress?
11.	How often do you use the internet to look at videos of cosmetic treatments, such as surgical procedures, related to your appearance concerns?
12.	How often do you use the internet for buying appearance related products due to your appearance concerns?
13.	How often do you use or visit web pages that purports to rate a person's attractiveness based on their appearance, to get feedback on your appearance concerns?
14.	How often do you use the internet, for example Google Search or social media, to find pictures or people to compare the parts of your appearance which cause you distress?
15.	How often do you avoid dating pages or dating apps, such as Tinder or Match.com, because of your appearance concerns?

Note. A 5-point Likert scale from never (1) to very often (5) was used for all items.

Table 3*Participants Demographics in sample 1*

Demographic	<i>M</i> (SD) or %
Age	33.13 years (11.34)
Gender	
Male	23.2 %
Female	75.5 %
Other	1.3 %
Nationality	
Sweden	97.3 %
Other	2.7 %
Work status	
Full time	41.9 %
Part time	7 %
Student	40.4 %
Sick leave	2.7 %
Unemployed	1.4 %
Other	6.6 %
Civic status	
Single	30.1 %
Married	23.5 %
Cohabitant	30.4 %
In a relationship	13 %
Other	3 %
Have been sick on leave due to appearance concerns?	
Yes, Currently	0.6 %
Yes, Before	2.7 %
Never	96.7 %
Sought treatment or care due to appearance concerns	
Surgical cosmetic	10.4 %
Psychological care	5.4 %
Medical care	13.3 %
Never	70.9 %

Statistical Analysis

IBM SPSS version 28.0.0.0 (190) was used to analyze the ARBO. A Principal Axis Factoring (PAF) Exploratory Factor Analysis was conducted using Promax rotation. Promax rotation is an oblique rotation method in other words, it assumes that the latent variables are correlated and rotates the data thereafter (DeVellis, 2017; Field, 2018; Knekta et al., 2019). This method of rotation was used since the potential latent variables of BDD and by extent the ARBO are

assumed to correlate. The descriptives, Kaiser-Meyer-Olkin values (KMO) and Bartlett's test of sphericity were checked to assess sample adequacy. The Anti-Imaging table was checked for covariance.

The number of latent variables included were determined using the Kaiser-Guttman Criterion (eigenvalues > 1), visual examination of the inflexion point of the scree-plot and Horn's Parallel analysis. Item deletion was determined to be a priority and items who did not load above .3 on any factor were deleted. Further, latent variables had to have a minimum of 3 items loaded upon them to be retained (DeVellis, 2017; Field, 2018; Knekta et al., 2019). McDonald's Omega was used to determine the current reliability of the ARBO and individual items were further removed if it would increase the reliability without damaging the theoretical framework of the ARBO. Further EFA's were conducted with a predetermined number of latent variables and items were deleted according to above stated criteria until a factor solution as close to optimal as possible, could be extracted.

Result 1

Exploratory Factor Analysis

1395 participants were included in the analysis, see Table 3 for demographics. An initial EFA was conducted on the 15 ARBO items shown in Table 1 using principal axis factoring and oblique rotation (promax). Bartlett's test of sphericity was significant, $\chi^2 (105, N = 1395) = 7221.45, p < .001$, and the KMO value was .87, indicating that the ARBO selected items were appropriate for factor analysis. The initial analysis was run to obtain eigenvalues for each factor in the data. Kaiser-Guttman criterion, scree plot and Horn's parallel analysis initially suggested a 3-factor solution. The extracted 3 factors accounted for 42.6% of the variance or 32.9%, 5.2% and 4.5% individually (corresponding eigenvalues of 5.45, 1.38 and 1.28). However, further examination of the items loading on each factor revealed that factor 3 consisted of only 2 items. This indicated that the 3-factor solution was not viable since one of

the factors had less than 3 items (Beavers et al, 2013; Field, 2018). Therefore, a second EFA was conducted using a predetermined two-factor solution with oblique rotation (promax). The 2 extracted factors accounted for 37.7% of the variance or 32.7% and 5% individually (corresponding eigenvalues of 5.45 and 1.37). A parallel analysis did not support the suggested factor structure, also item 3 & 15 had loadings below 0.3 on both factors. Furthermore, a reliability analysis on the suggested individual factor revealed that internal consistency measured by omega coefficient would increase if item 1, 3 & 15 were to be deleted. Thus, the mentioned items were deleted and a third EFA was submitted on the remaining 12 items with a two-factor solution using oblique rotation (promax). Barlett's test of sphericity was significant, $\chi^2 (66, N = 1395) = 6462.47, p < .001$, and the KMO value was .888. The extracted 2 factors accounted for 44.3% of the variance or 38.2% and 6.1% individually (corresponding eigenvalues of 5.1 and 1.25). After examining the Kaiser-Guttman criterion, scree plot, parallel analysis, items' individual eigenvalues and cross loadings, it was concluded that this solution was the best factor structure for the ARBO. The first factor included items 2 and 4-7 and is hypothesized to represent repetitive behaviors in an online environment and the second included item 8-14 hypothesized to be reassurance seeking in an online environment. Item loadings on the third model and previous ones can be seen in Table 4.

Table 4*Factor loadings of the three EFA models*

Items	Model 1			Model 2		Model 3	
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 1	Factor 2
Item 1			0.47	0.33			
Item 2	0.63			0.43		0.54	
Item 3			0.56	*			
Item 4	0.69			0.63		0.73	
Item 5	0.58			0.71		0.68	
Item 6	0.86			0.83		0.86	
Item 7	0.62			0.42		0.56	
Item 8		0.45			0.39		0.36
Item 9		0.56			0.75		0.63
Item 10		0.67			0.84		0.81
Item 11		0.70			0.76		0.78
Item 12		0.41			0.51		0.45
Item 13		0.58			0.57		0.54
Item 14		0.45			0.54		0.44
Item 15	0.40				*		

*Items that had < 0.3 loading on the factor and no higher loadings on other factors in the model.

Note. The third model the factor loadings of items 1, 3 & 15 are blank because they were deleted from the analysis.

Internal Consistency & Intercorrelations Between Factors 1

The internal consistency of the third model is excellent with McDonald's omega of .88.

Concerning the first factor, McDonald's omega is .83 and the factor/item correlations vary from .71 (Item 2) to .83 (Item 6), all significant, $p < 0.01$. Concerning the second factor, McDonald's omega is .82 and the factor/items correlations vary from .45 (Item 8) to .81 (Item 10), all significant, $p < 0.01$. Factor 1 and 2 scores are correlated at .62, $p < 0.01$.

Brief Discussion

A two-factor structure without items 1, 3 and 15 was extracted. Items 2 and 4-7 loaded onto factor 1 and item 8-14 loaded onto factor 2 with the first factor representing excessive repetitive behaviors in an online environment and the second factor representing reassurance seeking in an online environment. The two-factor structure was supported by the scree-plot, the eigenvalues and the parallel analysis upon removal of the items. A test of the models internal consistency displayed excellent reliability and the inter-item correlation was also good across both factors.

Method 2

Analysis two was conducted using data gathered between 2020-2022. The data was collected via snowball recruitment through various websites and SNS-groups in Sweden. An attempt to gather participants residing in Australia and Singapore was also conducted using the same test-battery. Those participants' nationality was labeled as *other* in Table 5. Being 18 years of age or older was once again listed as a requirement for participation and the total number of participants was 692. The full descriptives can be found in Table 5.

Table 5*Participants Demographics in sample 2*

Demographic	<i>M</i> (SD) or %
Age	28.09 years (9.93)
Gender	
Male	23.6 %
Female	74.6 %
Other	1.8 %
Nationality	
Sweden	88.5 %
Other	11.5 %
Work status	
Full time	32.8 %
Part time	5 %
Student	52.5 %
Unemployed	2.4 %
Civic status	
Single	41.8 %
Married	14.1 %
Cohabitant	20.8 %
In a relationship	22.1 %
Have been sick on leave due to appearance concerns?	
Yes, Currently	1.3 %
Yes, Before	11 %
Never	87.7 %
Sought treatment or care due to appearance concerns	
Surgical cosmetic	9 %
Psychological care	14.2 %
Medical care	14.4 %
Never	62.4 %

Measurements

Participants completed the ARBO in order to measure their online behavior. In addition, several other measures were included in the analysis to assess the convergent and divergent validity of the ARBO. *The Appearance and Anxiety Inventory* (AAI) (Veale et al., 2013) and the *Physical Appearance Comparison Scale–Revised* (PACS-R) (Schaefer & Thompson 2014) were used to specifically assess the convergent validity of the ARBO. *The Hospital Anxiety and Depression Scale* (HADS) (Zigmond & Snaith, 1983) was used for measuring anxiety symptoms, depressive symptoms and whether the ARBO diverges from those. *The Brief Obsessive Compulsive Scale* (BOCS) (Bejerot et al., 2014) and the *Body Appreciation Scale-2* (BAS-2) (Tylka & Wood-Barcalow, 2015) were also used to measure the divergent

validity of the ARBO. The scales were collected in a different sequence so that the order impacts could be examined. Four distinct versions of the questionnaire were produced as a consequence of the three rotations of the order. Since the order of the queries within each scale had to be preserved to guarantee that the scales were displayed as intended, full randomization was not a possibility. The sequence of the groups, the duration, and the total number of people for each rotation were recorded. No significant order effects were found across the groups (Barnes et al., 2020).

Convergent validity

The Appearance Anxiety Inventory (AAI). The AAI is a 10 item self-report scale aiming at measuring the cognitive processes behind BDD. It is based on trans-diagnostic assumptions of BDD stating that central themes of the disorder are distorted imagery and associated shame (Veale et al., 2014). The responses to these themes are cognitive processes and behaviors such as safety seeking which are common to BDD patients. The 10 items of the scale are scored on a 5-point Likert scale ranging from 0 = *not at all* to 4 = *all the time*. The scale has a maximum score of 40 (Veale et al., 2014). The AAI consists of a two-factor structure characterized by avoidance and threat monitoring. The AAI has demonstrated good convergent validity and the internal consistency of the scale has been deemed good (Veale et al., 2014). The AAI shares a trait with the ARBO in that it measures camouflaging as a form of avoidance (Veale et al., 2013). Similarly, the ARBO consists of items measuring image editing in order to hide perceived appearance flaws.

Physical Appearance Comparison Scale–Revised (PACS-R). The PACS-R is a self-report scale that consists of 11 items which were developed by Schaefer & Thompson (2014). This scale assesses the existence and degree of appearance comparison behaviors within both genders as opposed to the original Physical Appearance Comparison Scale (PACS) that only focused on women which was developed by Thompson et al. (1991).

Responses to the PACS-R are registered on a 5- point Likert scale (*0 = Never, 1 = Seldom, 2 = Sometimes, 3 = Often, 4 = Always*). The scale's predictive properties can capture aspects of body satisfaction and eating disorders. The psychometric properties of the scale have been examined and it has demonstrated satisfactory results (Schaefer & Thompson, 2014).

Divergent validity

Brief Obsessive–Compulsive Scale (BOCS). The BOCS is a self-report questionnaire that encompasses a total of 21 items which assess the symptoms and severity of obsessive-compulsive disorder in the DSM-V (Bejerot et al., 2014). The symptom subscale consists of 15 items that acquire information on whether the respondents engages in, have engaged or never engaged in 11 symptom areas (contamination, harming obsessions, sexual obsessions, checking, magical thoughts, morality, symmetry, counting, hoarding, somatic obsessions and self-harming behaviors). The severity subscale consists of 6 items reported on a 5-point Likert scale. The scale exhibits good specificity, sensitivity, internal consistency and inter-rater reliability (Bejerot et al., 2014).

The Hospital Anxiety and Depression Scale (HADS). The HADS comprises 14 items, 7 measuring anxiety symptoms and 7 measuring depression symptoms. Each item is scored on a 4-point Likert scale with several items being reverse coded. The score of the HADS ranges from 0-21. A score of 0-7 indicates no presence of anxiety or depression. A score of 8-10 indicates moderate symptoms of Anxiety and/or depression. And a score of 11 or higher indicates significant symptoms of either anxiety or depression. The HADS has been validated several times and have repeatedly been demonstrated to possess both good validity and good internal consistency in several contexts and on patients with varying diagnoses (Bjelland et al., 2002; Bocéréan & Dupret, 2014; Boxley et al., 2016; Wiriyakijja et al., 2020; & Cassiani-Miranda et al., 2021).

The Body Appreciation Scale (BAS-2). The BAS-2 is a scale measuring body appreciation or positive body image (Tylka & Wood-Barcalow, 2015). The BAS-2 is an updated version of the original BAS. It consists of 10 items loading on a unidimensional factor structure and has been adapted to disregard the sex of the participant. The items measure body appreciation on a five-point Likert scale with answer ranging from *Never* to *Always* (1-5) to questions such as *I feel love for my body* and *I feel like I am beautiful even if I am different from media images of attractive people* (Tylka & Wood-Barcalow, 2015). The BAS-2 have been shown to be internally consistent, possess good construct validity and the model-fit of the factor structure has been tested and deemed adequate (Tylka & Wood-Barcalow, 2015).

Ethical Considerations

The ethical considerations for study 2 were generally the same as for study one. The one difference being the data collection period of the second sample, which occurred between 2020 and 2022. The same test battery, which had previously been approved by Örebro University, was used for the second data collection. The same committee later approved the ethics of this second study before the datasets used in the study were handed out to the authors.

Procedure

Prior to the analysis participants who did not fully complete the questionnaire were labeled as missing and removed from the dataset utilizing listwise deletion. The same had to be done to participants who had answered alternative 6 on item 4. A total of 24 Participants were removed this way. A further 47 participants were removed as they were under the age of 18. The total number of participants in the second sample was 621. As was the case in study 1, no participants were removed from the second analysis due to nationality.

Statistical Analysis

A CFA was conducted using JASP v.0.17.1 based on maximum likelihood assumptions.

Model fit was used in order to confirm that the two-factor structure revealed by the EFA was applicable in identifying adults' appearance-related online behaviors, similar to those seen in BDD. Several different assessments were examined in order to check for model fit. The absolute fit index chi-square (χ^2) and the Standardized Mean Square Residual (SRMR) were used to analyze whether the predicted variance-covariance matrix correlates with the sample variance-covariance matrix and thus whether residuals in the model compromise the goodness-of-fit. The SRMR is based on standardized data and is generally considered to be more reliable than χ^2 but both are included in the results as the χ^2 is recommended by researchers. χ^2 is however vulnerable to the sample size of the model meaning that in samples with >200 participants χ^2 is almost always significant, indicating poor fit (Harrington, 2009). The Root Mean Square Error of Approximation (RMSEA) is used to balance out the χ^2 . The RMSEA is less sensitive to sample size but sensitive to model complexity. Finally, The Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) were used to evaluate if the model had good fit compared to a baseline model (Harrington, 2009).

For this analysis the following guidelines and cut-of values were used to assess goodness-of-fit. The χ^2 test indicates good model fit if the null-hypothesis is rejected ($p > 0.05$). The RMSEA should be 0.06 - 0.08 or less, suggesting reasonable error of approximation. The SRMR should be < 0.1 with more conservative researchers suggesting < 0.08 . CFI and TLI should be > 0.90 while others recommend close to or > 0.95 . It is worth noting that the TLI will generally be lower than the CFI (Harrington, 2009).

If the model was found to not fit well, certain changes were made if they could be properly motivated either theoretically or statistically, preferably both. The Residual covariance table was consulted and item pairs with modification indices (MI) of 100 or more

were further analyzed (Joachim et al., 2016). The Items were analyzed by reviewing relevant data on internet usage provided by Internetstiftelsen (2021) and common behaviors when measuring BDD. After reviewing an item pair with an MI of 100 or more the most problematic of the items were removed from the analysis and the analysis was then run again to check if the model-fit had improved. Revisions to the model using this method was done until the model-fit was deemed adequate according to above stated guidelines or theoretically justifiable modifications could no longer be found.

After the factor model had been modified to acquire adequate model fit, a further correlation analysis was conducted to examine its divergent and convergent validity. The reliability of the factor model was also examined using McDonald's Omega. Omega's reliance on a factor analytical approach makes it a more reliable option for assessing scales' internal consistency compared to the commonly used Chronbach's alpha (Trizano-Hermosilla & Alvarado, 2016). The HADS, BOCS and BAS-2 were used to analyze the factor models divergent validity while the PACS-R and the AAI were used to analyze the convergent validity of the factor model. Convergent validity in this case would indicate that the factor model is similar to other scales that measure appearance related behaviors or cognitions. Divergent validity would on the other hand indicate that the factor model differs from scales that measure different constructs such as anxiety, depression and OCD.

Result 2

CFA results

621 participants were included in the analysis. Results of the CFA using the factor structure indicated by the EFA, namely model 3, provided less than acceptable fit to the data, $\chi^2(53, N = 619) = 326,31, p < .001$, CFI = .9, TLI = .87, RMSEA = .09, SRMR = .06. Modification indices (MI) indicated residual covariances among pairs of items within the same subscale that shared an identical stem. The first item pair to be analyzed was item 5 and item 6 (MI =

102.92) *How often do you edit pictures of yourself to improve your appearance before you upload them to social media, because of distress related to your appearance concerns?* and *When you upload pictures of yourself to the internet, how often do you try to hide parts of your appearance due to distress related to your appearance concerns? (For example, editing pictures, using filters or by using accessories or decals to hide parts of your appearance?).*

The MI suggests that to reduce redundancy while maintaining adequate construct coverage one item from the pair may be eliminated from the subscale. Item 5 was deleted because it had a substantially lower factor loading as well as was deemed to be conceptually unfit as picture/image/video editing could be seen as another form of hiding aspects of one's appearance. Item 6 was thus deemed to encapsulate the theme of item 5. After item deletion a new CFA was conducted and the model fit was again analyzed. The new CFA indicated that the model fit had improved to an adequate level ($\chi^2(43, N = 619) = 212.3, p < .001$, CFI = .93, TLI = .91, RMSEA = .08, SRMR = .05). The summary of both models is provided in Table 6.

Table 6

Summary of KMO, Bartlett's test of sphericity and model-fit indices for the two models.

<i>Metric</i>	<i>Model 1</i>	<i>Model 2</i>
<i>KMO</i>	0.89	0.88
<i>Bartlett's test</i>	2726.19*	2342.82*
<i>CFI</i>	0.9	0.93
<i>TLI</i>	0.87	0.91
<i>RMSEA</i>	0.09	0.08
<i>SRMR</i>	0.06	0.05

Note. KMO, Kaiser–Meyer–Olkin test; CFI, Comparative Fit Index; TLI, Tucker–Lewis Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Residual.

* $p < 0.001$

Validity

The factor model appears to possess good convergent and divergent validity; more on the correlation between the factor model and its subscales (factor 1 and factor 2) and other measures, can be seen in Table 7.

Table 7

Summary of results from the tests of divergent and convergent validity on the ARBO.

<i>Scale</i>	<i>ARBO</i>	<i>ARBO subscale 1</i>	<i>ARBO subscale 2</i>
<i>HADS</i>	0.49	0.45	0.45
<i>HADS-A</i>	0.52	0.49	0.45
<i>HADS-D</i>	0.43	0.35	0.41
<i>BOCS</i>	0.38	0.35	0.35
<i>BAS-2</i>	-0.57	-0.57	-0.47
<i>PACS-R</i>	0.65	0.66	0.52
<i>AAI</i>	0.78	0.74	0.68

Note. HADS, Hospital Anxiety and Depression Scale; HADS-A, Hospital Anxiety and Depression Scale - Anxiety Subscale; HADS-D, Hospital Anxiety and Depression Scale - Depression Subscale; BOCS, Brief Obsessive Compulsive Scale; BAS-2, Body Appreciation Scale - 2nd version; PACS-R, Physical Appearance Comparison Scale - Revised; AAI, Appearance Anxiety Inventory. Table 7 only contains correlation values relevant to the aim of this thesis. All correlation coefficients included are significant, $p < 0.01$.

Internal Consistency & Intercorrelations Between Factors 2

The internal consistency of the final revised model (CFA - model 2) is excellent with McDonald's omega of .88. Concerning the first factor, McDonald's omega is .81 and the factor-item correlations vary from .74 (Item 2) to .83 (Item 6), all significant, $p < 0.01$. Concerning the second factor, McDonald's omega is .82 and the factor-item correlations vary from .51 (Item 8) to .81 (Item 9), all significant, $p < 0.01$.

Factor 1 and 2 scores are correlated at $.63, p < 0.01$. Additionally, the correlation between factor model and factor 1 is $.89, p < 0.001$ and $.92, p < 0.001$ between the factor model and factor 2. The internal consistency of the factor model thus appears to be good and the intercorrelations between the factors are significant.

Discussion

The aim of this study was to explore the psychometric properties of the ARBO self-report scale as well as exploring and later confirming a factor structure. Research has begun connecting PIU and appearance-related online behavior with BDD with a large literature review connecting them (Holland & Tiggemann, 2016). Common aspects of BDD such as repetitive behavior, information seeking and avoidance have further been hypothesized to manifest in an online environment (Ryding & Kuss, 2020). Due to the difficulties in diagnosing BDD, the possibility to screen for predictive online behaviors linked to BDD would be of great assistance to clinicians. The ARBO was developed in order to detect these appearance-related behaviors. The ARBO has previously been found to correlate well with several concurrent BDD scales and this thesis and its results builds upon the original work on the ARBO. This thesis is the next step towards making the ARBO a functional scale.

In study one, an EFA was used to explore the latent factor structure of the ARBO as well as delete items that did not fit the model. The analysis suggested that a two-factor structure was the best fit to the sample. Internal consistency was deemed excellent for the first sample. In study two, a CFA was conducted to confirm the model-fit of the explored factor structure as well as re-test the internal consistency and examine the convergent and divergent validity of the scale.

The two studies resulted in a two-factor structure believed to measure excessive repetitive behaviors and reassurance seeking in an online environment. The analyzed factor model on the two samples consists of 11 items with 4 loading on to the first factor and 7

loading on to the second factor. The model-fit of the factor model was deemed adequate and it demonstrated good convergent and divergent validity as well as excellent internal consistency. The results suggest that the 15 symptom-items of the ARBO consists of two factors. It also suggests that the now confirmed factor model measures what it intends to measure, appearance-related repetitive behaviors in an online environment.

Factor analysis

Prior to the analysis, several items were removed for a variety of reasons. These items included questions asking if the participants used social media or not as well as the amount of time spent engaged in the symptom behaviors and the rate of distress felt during said engagement. As previously stated, it was beyond the scope of this thesis to examine methods to include those items in the analysis as it would have meant redesigning several parts of the ARBO as well as potentially adding more items to the scale. This would not have been concurrent with factor analysis as a research method since one of its aims is to only retain relevant items and delete redundant ones. Further, several of the items that were removed from the analysis would have warranted their own latent variable consisting of only one or two items which are too few per latent variable (Field, 2018). However it is still important to understand the changes that have been made to the scale since the beginning of this thesis.

First of all, the removal of the original item 1 and 2 from the analysis is unlikely to have had an effect on the result. Participants who answered *no* on item 1 would have been labeled as missing prior to the analysis and then removed from the sample due to listwise deletion. Item 2 on the other hand was deemed to be similar to item 3 and thus excluded from the analysis. Item 18 which measured how time consuming the repetitive online behaviors were was also not included in the analysis. This is generally a good indicator of BDD since many patients report spending 3-8 hours a day engaged in appearance-related repetitive behaviors (Phillips, 2009 & APA, 2013). Further, time spent on SNS has been linked to

increased body surveillance and body-image concerns among adolescents and adults (Holland & Tiggeman, 2016; Vukovic et al., 2018). It has also been associated with both avoidance and distress among BDD patients (Phillips, 2014). However, this specific item alone has not been shown to correlate with any of the other symptom items. Examining this single item further was beyond the scope of this study. Similarly, the final items of the ARBO which measures distress were also removed prior to this analysis. While the level of distress is a good indicator of BDD and is a criterion for the disorder (APA, 2013) retaining the items in the analysis would have been problematic as it could potentially have meant a new latent variable with just two items. To summarize, the items removed prior to the analysis were removed for a variety of reasons. mainly because they represented additional constructs which were measured with too few items to warrant their own latent variable.

Several items were also removed during the EFA. Item 1 was removed to improve the internal consistency of the scale whilst Items 3 and 15 were removed in order to improve the internal consistency but also due to them loading too weakly on both factors. The two-factor structure achieved by the EFA ,which is an item reduction technique that aims to validate measurement scales and reduce total items if necessary, means that the items deleted during the EFA should be excluded from the scale.

Item 1 stated *How often do you post pictures of yourself on social media?*. This item was first and foremost deleted in order to improve the internal consistency but it also loaded poorly on both factors. The cut-off value when deciding whether an item loads on a factor was .3 for this study but it can also be .4 depending on whose criteria is used (Field, 2018). With item 1 only having a factor loading of .332 on factor 1; an argument could have been made here to remove it from the analysis but it was retained as the cut-off point for this analysis was .3. However, the decision to remove it was finally made in order to make the factor structure more reliable. Further, the removal of item 1 is supported by how the

inhabitants of Sweden use social media. There are many people who do not post images of themselves online at all as reported by Internetstiftelsen (2022). Only 25% of facebook users post original posts, pictures or videos on facebook. The same goes for other SNSs with more emphasis on image uploading such as instagram, Snapchat and TikTok. Out of instagram users, 46% upload their own images on the app. Out of Snapchat users, Only 25% upload images and videos via the sharing function *my story*. Finally, only 19% of TikTok users upload their own videos on the SNS (Internetstiftelsen, 2022). Depending on which SNS is used this means that 54-81% of users do not upload images. With a majority of people not uploading images of themselves on SNSs, it would be difficult to motivate the retention of the item even if it did not decrease the internal consistency of the factor structure.

Similarly, Item 3 was deleted due to its low loading on either of the two factors as well as to enhance the internal consistency of the subscale it loaded under (factor 1). The item is formulated as *How often do you post pictures of yourself on social media in order to get feedback regarding your appearance concerns? (For example, uploading pictures or writing posts about appearance)*. Such phrasing is problematic as it gives an example of an action (*writing posts*) that is totally different from the original action it is trying to measure (*frequency of uploading pictures*). Aside from the problematic wording and the low factor loading, there appears to be no correlation between getting appearance commentary and self-objectification. In a study by Burnell et al. (2021) analyzing a sample of young women using instagram the commentary received on posts was not correlated to body surveillance, body dissatisfaction, facial satisfaction or self-worth. This could very well indicate that people who suffer from appearance concerns or perhaps even BDD, generally do not seek feedback or post unfiltered images of themselves online. However, the question could be better if it were formulated differently. It could instead seek to measure the ratio between pictures posted on SNSs and the number of comments they receive as this was better

associated with higher levels of body surveillance and self-objectification (Burnell et al., 2021). The question could also simply be improved by changing the wording to *how often do you write posts or upload images on social media in order to get feedback regarding your appearance concerns*, though this is unlikely to make a difference as the focal point of the question remains the same.

Item number 15 that stated *How often do you avoid dating pages or dating apps, such as Tinder or Match.com, because of your appearance concerns?* was also deleted due to lower factor loading, as described earlier. Apart from that, it appears to lack theoretical generalizability and generated low variance during measurement. The item alludes to a section of the population that might use dating sites which are single people. Even with this specificity toward a subgroup of the total population, the item is not generalizable to all single people among which only 27% reported that they have used a dating site during the last year to go on a date (Internetstiftelsen, 2022). In the 2 samples used for this thesis, 30% and 41% respectively stated that their relationship status was *single*. This means that participants who are in a relationship are likely to disregard this question as they are unlikely to have a reason to use a dating site/app. However, it is still true that the number of people with BDD that are also single are higher than in the general population, 27% vs. 18% according to Brohede et al. (2015). Further, there are several reasons to avoid dating apps which could be attributed to *appearance concerns*. It is possible that people with anxiety disorders such as SAD or GAD and people with eating disorders would be avoiding dating sites for similar reasons. The item also refers to avoidance specifically due to appearance concerns but does allude to something akin to a non-behavior with its formulation. The formulation could perhaps be changed to *avoid due to appearance anxiety experienced when using dating apps* or *avoid dates or conversation on dating apps due to distress related to appearance concerns* in order to encapsulate something closer to a learned behavior or the

activity of actually using a dating app and not just the final product of going on a date. This item does have its merits as being single and avoiding relationships is common among BDD patients (Phillips, 2009 & Brohede et al., 2015) but the item itself excludes most participants, does not encapsulate the entire experience of using a dating app and is also not statistically viable in its current state.

In the original unpublished manuscript, The ARBO was not originally constructed with a theoretical factor structure (see Barnes et al., 2020). Rather, items were created by analyzing the most common behaviors among BDD patients and modifying them to an online environment (Phillips et al., 2006; Phillips, 2017). A 2-factor structure appears to be the current optimal solution for the samples used. The proposed factors, excessive repetitive behaviors and reassurance seeking, are currently constructs based on item descriptions. Although this does explain the conceptual overlap between certain items within the two factors, for example appearance comparison appears to be an aspect of both factors. The issue is lessened through the process of an EFA as well as tests of convergent and divergent validity but tests that compare the two latent variables with similar constructs will be necessary for future research on the ARBO.

Convergent & divergent validity

After establishing a factor structure with adequate model fit, five scales measuring relevant constructs were chosen to examine the factor model's convergent and divergent validity. To begin with, divergent validity was first assessed with the The HADS scale which showed a modest correlation to the factor model. This suggests that people that engaged in checking behaviors might exhibit anxiety and/or depressive symptoms. As it was evidenced by Koronczi et al. (2013), body dissatisfaction had a negative direct effect on problematic internet usage which was mediated by the effects of depressive and anxiety symptoms. Furthermore, it was also revealed that satisfaction with body image was both directly and

indirectly (via self-esteem and depression) associated negatively with problematic internet use (Koronczai et al., 2013). Further, women diagnosed with BDD reported higher levels of depression and anxiety as well as a higher risk of developing such symptomatology compared to the general population (Brohede et al., 2015). This suggests divergent validity between the HADS and the factor model while comorbidity is an explanation for the modest correlation between appearance-related behavior and depression and anxiety.

The BOCS was the second scale that was examined in relation to the factor model. As can be seen in Table 6, the BOCS exhibited low association with the factor model. The small correlation coefficient indicates that the factor model and BOCS measure different constructs, providing evidence for divergent validity. However, it is noticeable to see two constructs that share many similar aspects (repetitive behaviors and reassurance seeking are essential parts of OCD) have such a low association. This might be attributed to the wide range of obsessions that people with OCD might have (e.g. contamination, checking, harm, symmetry...etc) in contrast with what the factor model is measuring; body related concerns. Additionally, the ARBO is tailored to measure such behaviors in an online environment compared to BOCS that assess OCD behaviors mainly in offline settings. Finally, although there exists some comorbidity between the disorders, many people with BDD are misdiagnosed with OCD (Buckley et al., 2019). The results thus emphasize the factor model's usefulness as they indicate that it has the potential to differentiate between the two diagnoses.

BAS-2 was the third scale to be analyzed in relation to the factor model in order to examine divergent validity. The factor model showed negative correlation with BAS-2. This is explained by the fact that the two scale measure constructs of the opposite ends of the appearance appreciation spectrum. The ARBO measures ways to exhibit appearance concerns through online behaviors and the BAS-2 measures appearance appreciation. Previous examples of examining the relationship between appearance-preoccupation and

appearance-appreciation have shown similar results (Stojanovic & Tiveljung, 2022). In light of these results, it can be stated that the factor model is measuring constructs that are different from anxiety, depression, obsessive compulsive disorder that might be associated with appearance-concerns. Also, it shows that the factor model measures constructs opposite to body appreciation.

On the convergent validity front, the factor model showed an adequate correlation with the PACS-R. The correlation, however strong, does not suggest that the scales are measuring the same constructs. This points at the uniqueness of the factor model when compared to similar measurement scales. The factor model had a stronger correlation with the AAI scale despite not utilizing it during the construction process. This is indicative of the similarity of the constructs being measured (in other words, appearance anxiety and appearance-related behaviors that are caused by concerns and distress). The results of analyzing the convergent validity of the factor model indicate that the scale is similar to other scales measuring appearance-related constructs but does not suffer from collinearity with the AAI nor the PACS-R.

The confirmed factor structure captures PIU's multidimensionality coupled to appearance-related behaviors. It includes items measuring aspects of SNS but also expands on this by taking into account other aspects of appearance-related PIU such as cosmetic surgeries, appearance-related products and seeking reassurance online. This multidimensionality of the factor model might be the difference seen in convergent validity scales, meaning the extra aspects of the factor model made it not totally similar with PACS-R. The correlation between the factor model and PACS-R is strong but not above a threshold of 0.9 which would indicate that the two scales likely measure the same construct (Hair et al., 2019).

Given the factor model has demonstrated good convergent and divergent validity it is likely a suitable tool for screening tools used to understand appearance-related behavior and BDD online. The factor model possesses excellent convergent validity with concurrent scales measuring similar constructs but is also different enough from the other measures in this analysis to make it unique. As society moves further into the online realm, scales like the ARBO may become more important. The ARBO currently does a good job of filling the existing gap in scales measuring online appearance-related behaviors typical of BDD.

Limitations & Strengths

As with any thesis, there are multiple strengths and weaknesses that have to be taken into consideration. Selected limitations will be discussed first, followed by selected strengths.

Firstly, the samples used were not representative of a global or national population. For example, 40.4 and 52.5% of the participants identified themselves as students and a large majority of the participants identified as females (75.5 and 74.6 respectively). This imbalance among gender participation can potentially interfere with the results of the EFA, the CFA as well as the process of deleting and retaining items. Given that statistical tests reflect the variance within and between variables such as gender, having more representation of a certain gender will tilt the results towards it. In this case, the representation of females in the sample is higher than that of the general populations both in Sweden and globally - 49.5% and 49.7% respectively as of 2021 (World Bank, 2023; SCB, 2022). Consequently, this high skew towards females makes the results less generalizable to the rest of the population.

Although many of the ARBO items were not designed to be gender specific this could mean that this analysis of its psychometric properties is skewed and that the ARBO is more valid when measuring groups with a large portion of female participants. The issues arising from the gender proportions in the samples of this thesis is a result of the sampling method used

and can be addressed by future researchers by ensuring a sample which is representative of the population across the gender-spectrum.

Secondly, since the data used for the statistical analysis had already been collected when this thesis began the authors were limited in which scales could be used to analyze construct validity. This thesis was thus limited to assessing convergent and divergent validity using the available scales from the test-battery. Adding more scales to the test battery could have offered a more nuanced perspective of the ARBOs construct validity. An example of this is the previously mentioned SMAPS by Zimmer et al. (2021) which would have been a sound addition as it is the only other scale the authors of this thesis could find which measures appearance behaviors online. The authors of this thesis were constrained by the data available to them and future research on the ARBO and appearance related online behavior should be considering comparing the ARBO to relevant scales measuring similar or distinctly different constructs.

On the other hand, the thesis also possesses several strengths. Firstly, the statistical analysis used in this thesis generally utilizes best practice methods. EFA for example is a more accurate method for exploring the factor structure of a scale compared to the commonly used alternative, Principal Component Analysis (PCA) (Kahn, 2006; DeVellis, 2017). PCA aims to explain the maximum total variance and therefore assumes no error variance making it less preferable compared to an EFA, which only attempts to explain maximum common variance (Field, 2018). Another example is that the two studies utilize McDonald's Omega rather than Chronbach's Alpha for measuring internal consistency. Omega is a much more precise tool which is less vulnerable to several assumptions that can not be seen in data collected on larger samples (e.g. not all variables are going to be normally distributed). Additionally, this study used an oblique rotation method for the EFA. Oblique rotations assume that the latent variables are correlated and rotate them accordingly (DeVellis, 2017;

Field, 2018). Since the latent variables of the confirmed factor structure were found to be correlated, using a different rotation method, such as an orthogonal rotation, would have been erroneous. The mentioned methodologies and metrics are examples of how this thesis uses both an integrative theoretical and technical approach through the process of decision making. Multiple other factors and choices have been made utilizing this integrative approach, which is a strength in and of itself.

An additional strength of this thesis is conducting the EFA and CFA on two different samples. Majority of psychometric research conducts both analyses on one sample by randomly splitting it in two halves. This approach was avoided as multiple studies have shown its incapability of accurately assessing factor structure of questionnaires (Worthington & Whittaker, 2006; Cabrera-Nguyen, 2010; Boateng et al., 2018). During the preparation for this thesis, an initial decision of combining the two datasets was made as it appeared to be the normal approach in conducting psychometric research. This decision was later altered since a commitment to following best practices and recommendations was made. This finally led to conducting the EFA on one sample followed by CFA on the second one.

A final strength of the ARBO is that many of the items are formulated in a way that is inclusive across the gender spectrum. An example of this is the current item 12 (as it can be seen in Table 2) *How often do you use the internet for buying appearance related products due to your appearance concerns?*. By not mentioning specific appearance altering products the item encapsulates participants across the gender spectrum. This is very important since men and women with BDD often focus on different body-parts and thus may use different types of appearance related products. Men are more prone to focus on musculature while women are generally more focused on specific parts of their face, breasts, stomach etc (Phillips, 2009; APA, 2013). Phrasing the items to be more inclusive also decreases the amount of items needed for the scale, making it more efficient.

Finally, the ARBO is a self-report scale meaning it is vulnerable to the participants' subjective beliefs about their own behavior. All items included in the confirmed factor model are answered on a five-point Likert scale meaning that the participant answers the item according to their own subjective belief about their appearance related online behaviors. This subjectivity can be considered both a strength and a weakness of the ARBO.

It can be considered a strength as the participants' subjective suffering needs to be taken into account when examining BDD and appearance related behaviors. This is especially true when considering that a BDD diagnosis, like many other forms of mental illness, is based on a range of subjective experiences such as perceptions and distress (APA, 2013). The ARBOs Likert items are thus partly a strength, since the participants subjective belief is considered important when assessing BDD and appearance related behaviors.

The ARBOs Likert scale items may however also be limiting the scale as it may mean less objective answers. Participants may answer the items differently depending on their own subjective interpretation of the question and the answer. For example, the answer *very often* may mean different things to different people and may thus lead to different answers. The same is true for the subjective difference between the different answers, such as the difference between *seldom* and *sometimes* (option 2 and 3 respectively on the ARBOs five-point Likert scale). The subjective nature of Likert scale items may mean that two people with the same clinical distress and functional impairment give two different answers based on their interpretation of the item. This issue is partly negated by the fact that the ARBO possesses excellent internal consistency but Likert items are always going to be partly vulnerable to subjectivity and interpretation.

Future Research

Future research could cover areas like sampling and comparing between clinical and nonclinical populations. Additionally, it's beneficial while conducting such analyses to derive

a cut-off value that makes a distinction on the severity of the behaviors indulged in. On this front, it would be interesting to examine the predictive validity of the ARBO through a longitudinal analysis. Apart from that, further examination of the ARBO's properties such as sensitivity and specificity are beneficial and aid in garnering a better understanding of the latent variables that are being measured. Further, it would enable the ARBO to be used in a clinical setting with a lower risk for misdiagnosis. Additionally, future research should also aim towards comparing the ARBO with the SMAPS by Zimmer et al. (2021). The SMAPS measures online behaviors through SNS usage believed to be predictive of BDD and a side-by-side analysis of the two could provide both scales with incremental validity. This could improve both scales by highlighting overlap and differences. It could also give a clear image of the difference in analyzing appearance-related behaviors through PIU compared to only SNS usage.

Multiple items in the ARBO are not gender specific and analyzing the psychometrics of the ARBO with this in mind across different patient groups could be beneficial. One limitation of the ARBO's data is the age variant. A large portion of the participants in the two samples consisted of students with a mean age of around ~30. Further, a study by Enander et al. (2018) found that BDD prevalence, specifically for males, increased with age. Additionally, Internetstiftelsen (2021) reports that internet usage among older people has increased to the point where over 80% of people born as late as the 50's use the internet on a weekly basis. To analyze the ARBO on a specifically older population was not possible considering the available data. However, an analysis of the ARBOs psychometrics on an older population could make it more applicable to people regardless of age.

Acknowledgment

We would like to thank our supervisor John Barnes Ph.D. for providing us with the data sets used in this thesis. We would also like to thank him for his support and for sticking with us through the ups and downs of this thesis.

We hope you enjoyed reading and/or grading this thesis and wish you a good summer.

References

- AERA, APA, & NCME Joint Committee. (2014). *Standards for educational and psychological testing* (5th rev.). Washington: American Psychological Association.
- American Psychiatric Association (APA). (2013). *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDER* (5th ed.). American Psychiatric Publishing.
- Anastasi, A. (1985). Psychological testing: Basic concepts and common misconceptions. In A. M. Rogers & C. J. Scheirer (Eds.), *The G. Stanley Hall lecture series*, Vol. 5, pp. (87–120). American Psychological Association. <https://doi.org/10.1037/10052-003>
- Barnes, M., Abhyankar, P., Dimova, E., & Best, C. (2020). Associations between body dissatisfaction and self-reported anxiety and depression in otherwise healthy men: A systematic review and meta-analysis. *PloS one*, 15(2), e0229268. <https://doi.org/10.1371/journal.pone.0229268>
- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, 18(1), 6. <https://doi.org/10.7275/QV2Q-RK76>
- Bejerot, S., Edman, G., Anckarsäter, H., Berglund, G., Gillberg, C., Hofvander, B., Humble, M. B., Mörtberg, E., Råstam, M., Ståhlberg, O., & Frisén, L. (2014). The Brief Obsessive-Compulsive Scale (BOCS): a self-report scale for OCD and obsessive-compulsive related disorders. *Nordic journal of psychiatry*, 68(8), 549–559. <https://doi.org/10.3109/08039488.2014.884631>

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research*, 52(2), 69–77.

[https://doi.org/10.1016/S0022-3999\(01\)00296-3](https://doi.org/10.1016/S0022-3999(01)00296-3)

Boateng, Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, 6, 149–149.

<https://doi.org/10.3389/fpubh.2018.00149>

Bocéréan, C., & Dupret, E. (2014). A validation study of the Hospital Anxiety and Depression Scale (HADS) in a large sample of French employees. *BMC Psychiatry*, 14. <https://doi.org/10.1186/s12888-014-0354-0>

Boxley, L., Flaherty, J. M., Spencer, R. J., Drag, L. L., Pangilinan, P. H., & Bieliauskas, L. A., (2016) Reliability and factor structure of the Hospital Anxiety and Depression Scale in a polytrauma clinic. *JRRD*, 53(6), 873-880. doi: 10.1682/JRRD.2015.05.0088

Bridley, A., & Daffin L, W., (2022). *Fundamentals of psychology disorders (formerly abnormal psychology)* (3rd ed.). Washington State University.

Brohede, S., Wingren, G., Wijma, B., & Wijma, K. (2015). Prevalence of body dysmorphic disorder among Swedish women: a population-based study. *Comprehensive psychiatry*, 58, 108–115. <https://doi.org/10.1016/j.comppsy.2014.12.014>

Buchanan, B.G., Rossell, S.L., & Castle, D.J. (2011). Body dysmorphic disorder: a review of nosology, cognition and neurobiology. *Neuropsychiatry*, 1(1), 71–80.

Buckley, V., Krebs, G., Bowyer, L., Jassi, A., Goodman, R., Clark, B., & Stringaris, A. (2018). Innovations in Practice: Body dysmorphic disorder in youth - using the

Development and Well-Being Assessment as a tool to improve detection in routine clinical practice. *Child and adolescent mental health*, 23(3), 291–294.

<https://doi.org/10.1111/camh.12268>

Burnell, K., George, M. J., Kurup, A. R., & Underwood, M. K. (2021). "Ur a freakin goddess!": Examining Appearance Commentary on Instagram. *Psychology of popular media culture*, 10(4), 422–433. <https://doi.org/10.1037/ppm0000341>

Cabrera-Nguyen, P. (2010). Author Guidelines for Reporting Scale Development and Validation Results in the *Journal of the Society for Social Work and Research*. *Journal of the Society for Social Work and Research*, 1(2), 99–103.

<https://doi.org/10.5243/jsswr.2010.8>

Cash, T. F., Phillips, K. A., Santos, M. T., & Hrabosky, J. I. (2004). Measuring “negative body image”: Validation of the Body Image Disturbance Questionnaire in a nonclinical population. *Body Image: An International Journal of Research*, 1(4), 363–372. <https://doi.org/10.1016/j.bodyim.2004.10.001>.

Cassiani-Miranda, C. A., Scoppetta, O., & Cabanzo-Arenas, D. F. (2021). Validity of the hospital anxiety and depression scale (hads) in primary care patients in colombia. *General Hospital Psychiatry*. <https://doi.org/10.1016/j.genhosppsy.2021.01.014>

Clark, L. A., & Watson, D. I. (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment*, 7(3), 309–319.

<https://doi.org/10.1037/1040-3590.7.3.309>

- Crerand, Phillips, K. A., Menard, W., & Fay, C. (2005). Nonpsychiatric Medical Treatment of Body Dysmorphic Disorder. *Psychosomatics* (Washington, D.C.), 46(6), 549–555. <https://doi.org/10.1176/appi.psy.46.6.549>
- de Brito, M. J., Neto, M. S., de Oliveira, M. F., Cordás, T. A., Duarte, L. S., Rosella, M. F., Felix, G. A., & Ferreira, L. M. (2015). Yale-Brown Obsessive Compulsive Scale modified for Body Dysmorphic Disorder (BDD-YBOCS): Brazilian Portuguese translation, cultural adaptation and validation. *Revista Brasileira de Psiquiatria*, 37(4), 310–316. <https://doi-org.db.ub.oru.se/10.1590/1516-4446-2015-1664>.
- DeVellis, R. F. (2017). *Scale development : theory and applications (Fourth edition.)*. SAGE.
- Dima, A. L. (2018). Scale validation in applied health research: tutorial for a 6-step R-based psychometrics protocol. *Health Psychology and Behavioral Medicine*, 6(1), 136–161. <https://doi.org/10.1080/21642850.2018.1472602>
- Eisen, Phillips, K. A., Coles, M. E., & Rasmussen, S. A. (2004). Insight in obsessive compulsive disorder and body dysmorphic disorder. *Comprehensive Psychiatry*, 45(1), 10–15. <https://doi.org/10.1016/j.comppsy.2003.09.010>
- Enander, J., Volen, Z. I., Mataix-Cols, D., Kuja-Halkola, R., Ljótsson, B., Lundström, S., Pérez-Vigil, Ana., Monzani, B., Lichtenstein, P., & Rück, C. (2018). Prevalence and heritability of body dysmorphic symptoms in adolescents and young adults: a population-based nationwide twin study. *Psychological Medicine*, 48, 2740–2747. <https://doi.org/10.1017/S0033291718000375>.
- Etikprövningsmyndigheten. (2021, June 1). Stödmall forskningspersonsinformation. <https://etikprovningmyndigheten.se/wp-content/uploads/2021/06/Stodmall-forskning-spersonsinformation-Version-2021-06-01.docx>

Field. (2018). *Discovering statistics using IBM SPSS statistics (5.ed.)*. Sage Publications.

Gliem, J. A., & Gliem, R. R., Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Ohio: Ohio State University; 2003.

[accessed 2023 Feb 20]. Available at:

<https://scholarworks.iupui.edu/bitstream/handle/1805/344/gliem+&+gliem.pdf?sequence=1>

Hair, Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis (Eighth Edition)*. Cengage.

Harrington, D. (2009). *Confirmatory Factor Analysis*. New York: Oxford University Press.

Holland, & Tiggemann, M. (2016). A systematic review of the impact of the use of social networking sites on body image and disordered eating outcomes. *Body Image*, 17, 100–110. <https://doi.org/10.1016/j.bodyim.2016.02.008>

Internetstiftelsen. (2021). *Svenskarna och internet 2021*.

<https://svenskarnaochinternet.se/app/uploads/2021/09/internetstiftelsen-svenskarna-och-internet-2021.pdf>

Internetstiftelsen. (2022). *Svenskarna och internet 2022*.

<https://svenskarnaochinternet.se/app/uploads/2022/10/internetstiftelsen-svenskarna-och-internet-2022.pdf>

Joachim, K. C., Wilk, P., Ryan, B. L., & Speechley, K. N. (2016). Family-centered care in children with epilepsy: Evaluating the Measure of Processes of Care (MPOC-20).

Epilepsia, 57(10), 1660–1668. <https://doi.org/10.1111/epi.13494>

- Kazdin, A. E. (2016). Selecting measures for research investigations. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 205–226). American Psychological Association. <https://doi.org/10.1037/14805-013>
- Kahn, J. H. (2006). Factor Analysis in Counseling Psychology Research, Training, and Practice: Principles, Advances, and Applications. *The Counseling Psychologist*, 34(5), 684–718. <https://doi.org/10.1177/0011000006286347>
- Knekta, Runyon, C., & Eddy, S. (2019). One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. *CBE Life Sciences Education*, 18(1), rm1. <https://doi.org/10.1187/cbe.18-04-0064>
- Koronczai, B., Kökönyei, G., Urbán, R., Kun, B., Pápay, O., Nagygyörgy, K., Griffiths, M. D., & Demetrovics, Z. (2013). The mediating effect of self-esteem, depression and anxiety between satisfaction with body appearance and problematic internet use. *The American journal of drug and alcohol abuse*, 39(4), 259–265. <https://doi.org/10.3109/00952990.2013.803111>
- Kyriazos, T. A. (2018). Applied Psychometrics: The 3-Faced Construct Validation Method, a Routine for Evaluating a Factor Structure. *Psychology*, 9, 2044-2072. <https://doi.org/10.4236/psych.2018.98117>
- Littleton, H. L., Axsom, D., & Pury, C. L. (2005). Development of the body image concern inventory. *Behaviour Research and Therapy*, 43(2), 229–241. <https://doi.org/10.1016/j.brat.2003.12.006>
- Longley, S. L., Holm-Denoma, J., Allan, N. P., Calamari, J. E., Armstrong, K., Wainwright, A., & Hasan, N. (2019). A quantitative study of body dysmorphic disorder: Latent

- structure and correlates. *Journal of Obsessive-Compulsive and Related Disorders.*, 21, 82–90. <https://doi.org/10.1016/j.jocrd.2018.12.004>
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000144>
- Melioli, Rodgers, R. F., Rodrigues, M., & Chabrol, H. (2015). The Role of Body Image in the Relationship Between Internet Use and Bulimic Symptoms: Three Theoretical Frameworks. *Cyberpsychology, Behavior and Social Networking*, 18(11), 682–686. <https://doi.org/10.1089/cyber.2015.0154>
- Möllman, A., Dietel, F. A., Hunger, A., & Buhlmann, U., (2017) Prevalence of body dysmorphic disorder and associated features in German adolescents: A self-report survey. *Psychiatry Research*, 254 263–267. <http://dx.doi.org/10.1016/j.psychres.2017.04.063>
- Oosthuizen, P., Lambert, T., & Castle, D. J. (1998). Dysmorphic concern: Prevalence and associations with clinical variables. *Australian & New Zealand Journal of Psychiatry*, 32(1), 129–132. <https://doi.org/10.3109/00048679809062719>
- Oshana, A., Klimek, P., & Blashill A. J., (2020). Minority stress and body dysmorphic disorder symptoms among sexual minority adolescents and adult men. *Body Image*, 34(2020), 167-172. <https://doi.org/10.1016/j.bodyim.2020.06.001>
- Patel, T. A., Stewart, R. A., Summers, B. J., Wilver, N, L., & Cougle, J. R., (2022). A self-report assessment of appearance-related safety behaviours: Development and psychometric properties. *Behavioural and Cognitive Psychotherapy.*, 50(6), 604–619. <https://doi.org/10.1017/S135246582200039X>
- Petersson, E., & Tuupanen, M. (2020). Body dysmorphic disorder in a Swedish sample:

Prevalence and online behavior [Masteruppsats]. Örebro Universitet.

Phillips, K. A., Atala, K. D., & Pope, H. G. (1995). Diagnostic instruments for body dysmorphic disorder. New research, programs and abstracts. Miami, FL.

Phillips, K. A., (2004a). Body dysmorphic disorder: recognizing and treating imagined ugliness. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 3(1), 12–17.

Phillips, K. A., (2004b). Psychosis in body dysmorphic disorder. *Journal of Psychiatric Research.*, 38(1), 63–72. [https://doi.org/10.1016/s0022-3956\(03\)00098-0](https://doi.org/10.1016/s0022-3956(03)00098-0)

Phillips, K. A., & Menard, W., & Fay, C., (2006). Gender similarities and differences in 200 individuals with body dysmorphic disorder. *Comprehensive Psychiatry.*, 47(2), 77–87. <https://doi.org/10.1016/j.comppsy.2005.07.002>

Phillips, K. A. (2009). *Understanding body dysmorphic disorder: An essential guide*. Oxford University Press.

Phillips, K. A., (2017). Body dysmorphic disorder: Advances in research and clinical practice. *Oxford University Press*.
<https://doi.org/10.1093/med/9780190254131.001.0001>

Rousseau, Eggermont, S., & Frison, E. (2017). The reciprocal and indirect relationships between passive Facebook use, comparison on Facebook, and adolescents' body dissatisfaction. *Computers in Human Behavior*, 73, 336–344.
<https://doi.org/10.1016/j.chb.2017.03.056>

Rust, J., Kosinski, M., & Stillwell, D. (2020). *Modern Psychometrics: The Science of Psychological Assessment* (4th ed.). Routledge.

- Ryding, F. C., & Kuss, D. J. (2020). The use of social networking sites, body image dissatisfaction, and body dysmorphic disorder: A systematic review of psychological research. *Psychology of Popular Media*, 9(4), 412–435.
<https://doi.org/10.1037/ppm0000264>
- Statistiska Centralbyrån (SCB). (2022). På tal om kvinnor och män 2022: Lathund om jämställdhet.
https://www.scb.se/contentassets/b3ba3d3ad7a74749936c7fd2e3b4bee6/le0201_2021_b22_x10br2201.pdf
- Statistiska Centralbyrån (SCB). (2022). *Utrikes födda i Sverige*.
<https://www.scb.se/hitta-statistik/sverige-i-siffror/manniskorna-i-sverige/utrikes-fodda-i-sverige/>
- Schaefer, L. M., & Thompson, J. K. (2014). The development and validation of the Physical Appearance Comparison Scale-Revised (PACS-R). *Eating behaviors*, 15(2), 209–217.
<https://doi.org/10.1016/j.eatbeh.2014.01.001>
- Schieber, K., Kollei, I., de Zwaan, M., & Martin, A. (2015). Classification of body dysmorphic disorder — What is the advantage of the new DSM-5 criteria?. *Journal of Psychosomatic Research*, 78(3), 223–227.
<https://doi.org/10.1016/j.jpsychores.2015.01.002>
- Shapira, N. A., Goldsmith, T. D., Keck, P. E., Jr., Khosla, U. M., & McElroy, S. L. (2000). Psychiatric features of individuals with problematic Internet use. *Journal of Affective Disorders*, 57(1–3), 267–272. [https://doi.org/10.1016/S0165-0327\(99\)00107-X](https://doi.org/10.1016/S0165-0327(99)00107-X)
- Simbar, Nazarpour, S., Majd, H. A., Andarvar, K. D., Torkamani, Z. J., & Rahnemaei, F. A. (2020). Correction to: Is body image a predictor of women’s depression and anxiety in

postmenopausal women? *BMC Psychiatry*, 20(1), 278–278.

<https://doi.org/10.1186/s12888-020-02685-y>

Singh & Veale, D. (2019). Understanding and treating body dysmorphic disorder. *Indian Journal of Psychiatry*, 61(7), 131–135.

https://doi.org/10.4103/psychiatry.IndianJPsychiatry_528_18

Stojanovic, L.P., & Tiveljung, R. (2022). The Relation Between Body Appreciation and Appearance-Comparative Behaviors Online - Is It Moderated by Age? [Masteruppsats]. Örebro Universitet.

Streiner, Norman, G. R., & Cairney, J. (2015). *Health measurement scales : a practical guide to their development and use (5. ed.)*. Oxford University Press.

Thompson, J. K., Heinberg, L., & Tantleff, S. (1991). The Physical Appearance Comparison Scale (PACS). *The Behavior Therapist*, 14, 174.

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in psychology*, 7, 769. <https://doi.org/10.3389/fpsyg.2016.00769>

Tylka, T. L., & Wood-Barcalow, N. L., (2015). The Body Appreciation Scale-2: item refinement and psychometric evaluation. *Body Image*, 12, 53–67.
<https://doi.org/10.1016/j.bodyim.2014.09.006>

van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Sage Publications, Inc.

Veale, Eshkevari, E., Kanakam, N., Ellison, N., Costa, A., & Werner, T. (2014). The Appearance Anxiety Inventory: Validation of a Process Measure in the Treatment of

- Body Dysmorphic Disorder. *Behavioural and Cognitive Psychotherapy*, 42(5), 605–616. <https://doi.org/10.1017/S1352465813000556>
- Veale, D., Gledhill, L. J., Christodoulou, P., & Hodsoll, J. (2016). Body dysmorphic disorder in different settings: A systematic review and estimated weighted prevalence. *Body Image*, 18, 168–186. <https://doi.org/10.1016/j.bodyim.2016.07.003>
- Wiriyakijja, P., Porter, S., Fedele, S., Hodgson, T., McMillan, R., Shephard, M., & Ni Riordain, R. (2020). Validation of the HADS and PSS-10 and a cross-sectional study of psychological status in patients with recurrent aphthous stomatitis. *Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology*, 49(3), 260–270. <https://doi.org/10.1111/jop.12991>
- World Bank. (2023). *World Bank Open Data : Population, female (% of total population)*. <https://data.worldbank.org/indicator/SP.POP.TOTL.FE.ZS>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist*, 34(6), 806–838. <https://doi.org/10.1177/0011000006288127>
- Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica Scandinavica*, 67(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>
- Zimmer-Gembeck, M. J., Hawes, T., & Pariz, J. (2021). A closer look at appearance and social media: Measuring activity, self-presentation, and social comparison and their associations with emotional adjustment. *Psychology of Popular Media*, 10(1), 74–86. <https://doi.org/10.1037/ppm0000277>.